*The Journal of Laryngology & Otology (2016)*, ***130***, *104–106*.
© JLO (1984) Limited, 2015
doi:10.1017/S0022215115003151

**SHORT COMMUNICATION**

# Smartphone speech-to-text applications for communication with profoundly deaf patients

F C LYALL[1], P J CLAMP[2], D HAJIOFF[2]

[1]*Peninsula Radiology Academy, Plymouth, and* [2]*Department of ENT Surgery, University Hospitals Bristol, UK*

## Abstract

*Objective*: Visual communication aids, such as handwriting or typing, are often used to communicate with deaf patients in the clinic. This study aimed to establish the feasibility of communicating through smartphone speech recognition software compared with writing or typing.

*Method*: Thirty doctors and medical students were timed writing, typing and dictating a standard set of six sentences appropriate for a post-operative consultation, and the results were assessed for accuracy and legibility.

*Results*: The mean time for smartphone dictation (17.8 seconds, 95 per cent confidence interval = 17.0–18.7) was significantly faster than writing (59.2 seconds, 95 per cent confidence interval = 56.6–61.7) or typing (44 seconds, 95 per cent confidence interval = 41.0–47.1) ($p < 0.001$). Speech recognition was slightly less accurate, but accuracy increased with time spent dictating.

*Conclusion*: Smartphone dictation is a feasible alternative to typing and handwriting. Slow speech may improve accuracy. Early clinical experience has been promising.

**Key words:** Cell Phones; Communication Aids for Disabled; Communication Methods, Total; Cochlear Implantation; Persons With Hearing Impairments; Postoperative Care; Speech Recognition Software; Writing

## Introduction

Effective communication with patients with severe hearing loss can pose a challenge in healthcare settings. Even when healthcare professionals use 'deaf-aware' communication strategies to aid lip-reading, studies suggest that patients still miss essential information during the consultation.[1,2] Patients with congenital deafness tend to prefer using sign language, but those with progressive or recent hearing loss may not have this skill.[2,3] Communication via writing can be useful and is the most common technique employed by doctors.[1]

Patients with severe to profound hearing loss may benefit from cochlear implants, especially when external hearing aids fail to provide adequate benefit. Assessing patients for cochlear implantation requires a detailed evaluation of the hearing impairment, medical history, and psychological and social factors. Patients are counselled and educated about the risks and benefits, and post-operative rehabilitation. Effective communication is required at all stages of the process. In the UK, cochlear implants in adults are approved for profoundly deaf patients (with a minimum of 90 dB HL loss at 2 and 4 kHz) with aided speech recognition (Bamford–Kowal–Bench test) scores of less than 50 per cent at 70 dB SPL.[4] By definition, these patients find verbal communication difficult, even with well-fitted and tuned acoustic hearing aids.

In our clinics, we use a range of communication aids, including sign language interpreters and visual aids such as written text. Many of these patients are post-lingually deaf and do not use sign language. When written text is required, we favour typing conversations onto a desktop computer using word-processing applications. This allows clear text to be displayed on a monitor, and the font size can be increased as required.

In some situations, access to interpreters or computers is not possible. This includes many clinical areas such as emergency departments, wards, operating theatres and post-operative recovery areas. In these situations, we have had anecdotal success using automated speech recognition software on smartphones. This allows users to dictate a conversation directly onto the smartphone. Using mobile data communication, this is quickly converted to text that can be read from the screen. In order to more formally assess the feasibility of this technology, we conducted a study of its speed and accuracy compared with typing and writing.

## Materials and methods

In this prospective study, volunteer medical students and doctors were enrolled. These subjects were given a standard set of six sentences which reflected a typical conversation in the recovery room or ward following cochlear implant surgery (Appendix 1). The subjects were asked to replicate this text by writing the information down using pen and paper, typing it into a computer word processor, and dictating it into a smartphone running an automated speech recognition software application. For computer typing, we used Word 2003 software (Microsoft, Redmond, Washington,

Presented at the Association of Surgeons in Training conference, 5–7 April 2013, Manchester, UK.
Accepted for publication 23 September 2015 First published online 27 November 2015

https://doi.org/10.1017/S0022215115003151 Published online by Cambridge University Press

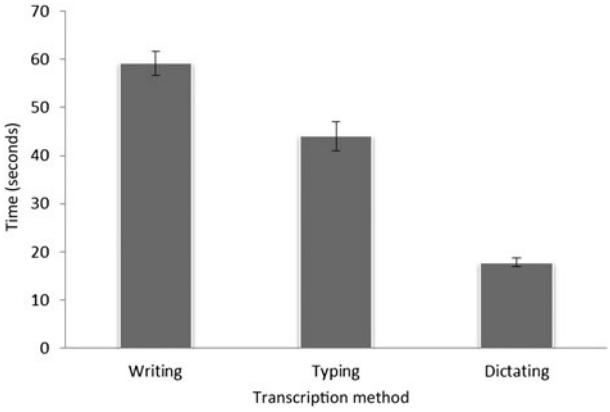



FIG. 1
Mean time taken for each transcription method (vertical lines represent 95 per cent confidence intervals).

USA). Automated speech recognition was performed using the free Dragon Dictation application (Nuance, Burlington, Massachusetts, USA) running on an Apple iPhone 4 (Apple, Cupertino, California, USA). The iPhone had third generation ('3G') network connectivity.

Each method of transcription was timed and the final result measured for accuracy. Timing did not include turning on the computer and logging on as it would normally be on for the duration of the clinic; however, it did include unlocking the computer and opening Word. Similarly, the timing of the dictation included unlocking the iPhone and opening the application. Handwriting was assessed for legibility by two independent readers. The accuracy of the typed and dictated information was scored in terms of the percentage of words correct.

## Results

Thirty subjects, 11 doctors and 19 medical students, were recruited.

The mean times and accuracy values for the three transcription methods are shown in Figures 1 and 2. The time and accuracy for each method were compared using the

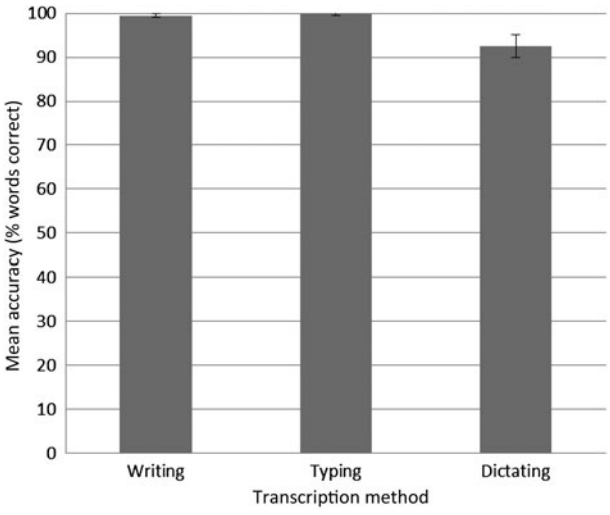



FIG. 2
Mean accuracy for each transcription method (vertical lines represent 95 per cent confidence intervals).

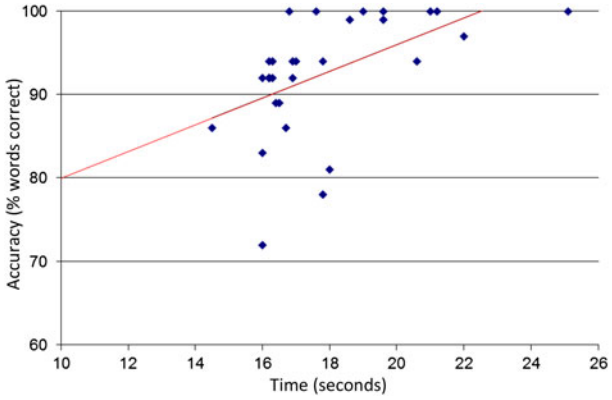



FIG. 3
Correlation between time taken to dictate and accuracy of dictation.

paired samples *t*-test (SPSS Statistics software, version 19; IBM, Armonk, New York, USA).

The mean time for smartphone dictation (17.8 seconds, 95 per cent confidence interval (CI) = 17.0–18.7) was significantly faster than writing (59.2 seconds, 95 per cent CI = 56.6–61.7) or typing (44 seconds, 95 per cent CI = 41.0–47.1) ($p < 0.001$). Smartphone dictation was less accurate than the other methods (dictation – 92.5 per cent words correct, 95 per cent CI = 89.8–95.1; writing – 99.5 per cent words correct, 95 per cent CI = 99.1–99.9; and typing – 99.8 per cent words correct, 95 per cent CI = 99.5–100.1) ($p < 0.001$).

### *Digital dictation*

The accuracy of smartphone dictation appeared to improve with time spent dictating (Spearman's rank correlation coefficient = 0.655; $p < 0.01$) (Figure 3).

### *Students versus doctors*

Students were faster than doctors at digital dictation (mean time of 16.7 *vs* 19.7 seconds respectively, $p < 0.001$), but less accurate (90.3 *vs* 96.4 per cent words correct respectively; $p = 0.004$ Mann–Whitney U test).

## Discussion

This pilot study is the first to show that the use of smartphone automated speech recognition software may be feasible in a clinical setting.

A review of published studies identified five papers that examined the use of visual text communication methods in the clinical setting.[5–9] Hinman *et al*. showed that comprehension was significantly better in deaf patients who read text generated by automatic speech recognition software (alone, with listening, or with listening and lip-reading) compared to listening alone.[5] However, this was a small study of eight patients with a range of mild to profound hearing losses. These findings have since been confirmed by Zekveld *et al*. in a study in which subjects with no hearing impairment demonstrated better speech comprehension against background noise when speech was accompanied by text generated by automated speech recognition software.[6]

Three papers have described the use of automated speech recognition software for administration in out-patient clinics. Klann and Szolovits used Dragon Naturally Speaking automated speech recognition software (Nuance) to produce consultation records.[7] However, the software worked better

https://doi.org/10.1017/S0022215115003151 Published online by Cambridge University Press

when the user wore a headset, which may not be acceptable during a consultation. One group studied the use of automated speech recognition software for creating paediatric out-patient clinic letters.[8] They found that training the software and correcting errors took double the physician time in comparison to using a handheld tape recorder, although the letters were available the same day, in comparison to one week using the standard practice. Bottrill described his use of the same software for producing ENT out-patient letters.[9] He found it freed up his secretary's time to check the letters for accuracy, and sped up the turnaround time; however, it took time to train the system.

We did not find any studies of automated speech recognition software used to communicate with patients directly, although the possibility has been described.[10]

Our study shows that smartphone automated speech recognition is significantly faster than writing or typing. Although the results for dictation were not quite as accurate, the error rate was only 7 per cent on average. In our experience with this technology, patients experience little difficulty understanding text with this low frequency of errors. There is significant redundancy in most language, which means that patients can understand the meaning of sentences even when some words are incorrect. Several studies have shown this to be the case, particularly for deaf patients.[5,6] Accuracy can be improved by slower dictation.

There are factors that may limit the use of this technology in clinical settings. We used the free Dragon Dictation application for the iPhone. This requires a connection to the internet via third generation, fourth generation ('4G') or Wi-Fi™ to process speech. Not all clinical areas have cellular reception, and many institutions do not provide wireless internet access for staff. The benefit of uploading the speech to the network is that the application learns different accents and vocabulary, and continuously improves. We anticipate that accuracy will increase as the technology improves. Another drawback is that the size of smartphone display limits the size of text displayed. Patients with poor eyesight may be better served by tablet computers.

Despite these limitations, we have found this technology to be a useful aid in communicating with patients in the out-patient clinic and recovery ward.

## Acknowledgements

Thank you to the medical students and doctors who gave their time to be subjects.

## References

1 Ebert DA, Heckerling PS. Communication with deaf patients: knowledge, beliefs and practices of physicians. *JAMA* 1995; **273**:227–9
2 Reeves D, Kokoruwe B. Communication and communication support in primary care: a survey of deaf patients. *Audiol Med* 2005;**3**:95–107
3 Middleton A, Turner GH, Bitner-Glindzicz M, Lewis P, Richards M, Clarke A *et al*. Preferences for communication in clinic from deaf people: a cross-sectional study. *J Eval Clin Prac* 2010;**16**:811–17
4 National Institute for Health and Clinical Excellence. *NICE Technology Appraisal Guidance 166: Cochlear Implants for Children and Adults with Severe to Profound Deafness*. London: NICE, 2009
5 Hinman RT, Lupton EC, Leeb SB, Avestruz A, Gilmore R, Paul D *et al*. Using talking lights illumination-based communication networks to enhance word comprehension by people who are deaf or hard of hearing. *Am J Audiol* 2003;**12**:17–22
6 Zekveld AA, Kramer SE, Kessens JM, Vlaming MS, Houtgast T. The benefit obtained from visually displayed text from an automatic speech recognizer during listening to speech presented in noise. *Ear Hear* 2008;**29**:838–52
7 Klann JG, Szolovits P. An intelligent listening framework for capturing encounter notes from a doctor-patient dialog. *BMC Med Inform Decis Mak* 2009;**9**(suppl 1):S3
8 Issenmann RM, Haffer IH. Use of voice recognition software in an outpatient pediatric specialty practice. *Pediatrics* 2004;**114**: 290–3
9 Bottrill ID. The practical application of voice recognition technology in the ENT clinic. *Otorhinolaryngologist* 2008;**2**:58–60
10 Fellbaum K, Koroupetroglou G. Principles of electronic speech processing with applications for people with disabilities. *Technology and Disability* 2008;**20**:55–85

## Appendix 1. Post-operative questions and information giving

How do you feel?
The operation went very well.
Do you have any pain? Do you feel dizzy or sick?
Can you close your eyes tightly and show me your teeth?
Do you have any questions?

Address for correspondence:
Dr Fiona Lyall,
Peninsula Radiology Academy,
William Prance Road,
Derriford,
Plymouth PL6 5WR, UK

E-mail: fiona.lyall@nhs.net

Dr F Lyall takes responsibility for the integrity of the content of the paper
Competing interests: None declared

https://doi.org/10.1017/S0022215115003151 Published online by Cambridge University Press